

The conditional entropy power inequality for Gaussian quantum states

Robert Koenig

April 29, 2013

Abstract

We propose a generalization of the quantum entropy power inequality involving conditional entropies. For the special case of Gaussian states, we give a proof based on perturbation theory for symplectic spectra. We discuss some implications for entanglement-assisted classical communication over additive bosonic noise channels.

1 Classical and quantum entropy-power inequalities

The entropy power inequality, proposed by Shannon [27] and later established with increasing rigor by Stam [29] and Blachman [5], has become a fundamental tool in classical information theory. Shannon's original application of the entropy power inequality is a lower bound on the capacity of an additive (but potentially non-Gaussian) noise channel [27, Theorem 18]. However, the usefulness of the entropy power inequality is especially evident in multi-terminal information theory. Among the most well-known applications are the characterization of the Gaussian broadcast channel by Bergman [4], the Gaussian two-description problem by Ozarow [25] and the quadratic Gaussian CEO problem by Oohama [24]. In these multi-user settings, Fano's inequality by itself is insufficient to characterize different tradeoffs. A more recent application proposed by Liu and Viswanath [23] uses entropy power inequalities to solve certain optimization problems.

The entropy power inequality lower bounds the differential entropy of the convolution of two independent random variables X, Y taking values in \mathbb{R}^n . Its covariance-preserving version states that

$$H(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \geq \lambda H(X) + (1-\lambda)H(Y) \quad \text{for all } 0 \leq \lambda \leq 1. \quad (1)$$

Eq. (1) can be shown [22, 32] to be equivalent to the more commonly used statement

$$e^{2H(X+Y)/n} \geq e^{2H(X)/n} + e^{2H(Y)/n}.$$

The latter explains the terminology as $e^{2H(X)/n}$ is the power, i.e., variance of a Gaussian random variable with identical entropy as $H(X)$. Inequalities such as (1) are closely related to Log-Sobolev inequalities (see e.g., [31]) as well as Brunn-Minkowski-type inequalities [6]. Generalizations to free probability [30] and quantum states have been considered.

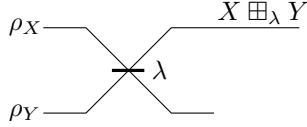


Figure 1: The setting of the (unconditional) quantum entropy power inequality: two independent sets of modes (i.e., a product state $\rho_X \otimes \rho_Y$) combine at a beam-splitter of transmissivity $\lambda \in [0, 1]$.

Among known quantum generalizations is the photon-number-inequality conjectured in [13] (and proved for special cases such as Gaussian states in [12]), and a version of (1) involving von Neumann entropies $S(\rho) = -\text{tr}(\rho \log \rho)$ instead of differential (Shannon) entropies [20]. Formally, this statement is obtained by substituting states ρ_X, ρ_Y of n bosonic modes for the random variables (X, Y) , and using a beamsplitter to process the product state $\rho_X \otimes \rho_Y$, see Fig. 1. This results in an output denoted $\rho_{X \boxplus_\lambda Y}$ on one of the output arms (see below for a precise definition), and the corresponding quantum generalization states that

$$S(X \boxplus_\lambda Y) \geq \lambda S(X) + (1 - \lambda)S(Y) \quad \text{for all } 0 \leq \lambda \leq 1. \quad (2)$$

As discussed in [21], Eq. (2) provides strong upper limits on classical communication over additive thermal noise channels. Indeed, for suitable parameter regimes, the resulting upper bounds on the classical capacity are close to the Holevo-Schumacher-Westmoreland lower bound achievable by coherent states. This limits the degree of potential additivity violations, implying that coding strategies using simple product states are close to optimal for thermal noise channels.

2 Conditional entropy-power inequalities

Most applications of the classical entropy power inequality make use of a version for conditional entropies $H(X|Z) = H(XZ) - H(Z)$. It is easy to see that, if (X, Y) are conditionally independent given Z , then

$$H(\sqrt{\lambda}X + \sqrt{1 - \lambda}Y|Z) \geq \lambda H(X|Z) + (1 - \lambda)H(Y|Z) \quad \text{for all } 0 \leq \lambda \leq 1. \quad (3)$$

Indeed, (3) is an immediate consequence of (1) and the fact that the conditional entropy $H(X|Z) = \sum_z P_Z(z)H(X|Z = z)$ is the average of the entropies $H(X|Z = z)$ of the conditional distributions $\{P_{X|Z=z}\}_z$.

It is natural to ask whether a generalization of (3) is true for states ρ_{XYZ} for which the n -mode systems X and Y are conditionally independent given the quantum system Z , i.e., whether

$$S(X \boxplus_\lambda Y|Z) \geq \lambda S(X|Z) + (1 - \lambda)S(Y|Z). \quad (4)$$

We will argue that (4) is useful to estimate entanglement-assisted capacities of additive noise channels. The inequality may have additional applications in multi-user quantum information theory.

Establishing an inequality of the form (4) appears to be non-trivial because we cannot simply condition on the quantum system Z (unless, of course, it is purely classical). However,

the following simplification is immediate: it suffices to establish an inequality of the form

$$S(X \boxplus_\lambda Y | Z_1 Z_2) \geq \lambda S(X | Z_1) + (1 - \lambda) S(Y | Z_2) \quad \text{for all product states } \rho_{XZ_1} \otimes \rho_{YZ_2} . \quad (5)$$

This is because any conditionally independent state ρ_{XYZ} has the Markov form [15]

$$\rho_{XYZ} = \bigoplus_j p_j \rho_{XZ_1^{(j)}} \otimes \rho_{YZ_2^{(j)}}$$

and the von Neumann entropy satisfies $S(\bigoplus_j p_j \rho_j) = \sum_j p_j S(\rho_j) + H(p)$ on direct sums, where $H(p)$ is the Shannon entropy of the distribution $\{p_j\}_j$.

Here we prove inequality (5) for all pairs of Gaussian states ρ_{XZ_1} and ρ_{YZ_2} . We conjecture that Eq. (4) holds in general for arbitrary conditionally independent states $\rho_{X_1 X_2 Z}$. If this is the case, the implications for entanglement-assisted capacities discussed here extend to all additive (but not necessarily Gaussian) channels.

It may be possible to find a proof of Eq. (4) using a similar strategy as we employ in the Gaussian case. However, this will require a novel analysis of the evolution of conditional entropies under a certain Markovian evolution. In the Gaussian case, this is based on perturbation theory for symplectic eigenvalues, as we explain below.

3 Implications for entanglement-assisted communication

Quantum communication channels are characterized by different capacities depending on the additional auxiliary resources available, as well as whether or not the communicated information is classical or quantum. Here we focus on entanglement-assisted classical capacities which are arguably best understood. Consider a point-to-point scenario where a sender A tries to communicate to a receiver C over a channel $\mathcal{E} : \mathcal{B}(A) \rightarrow \mathcal{B}(C)$. The entanglement-assisted classical capacity $C_E(\mathcal{E})$ is defined as the maximal rate (in bits/channel use) at which classical bits can be transmitted reliably if the sender and receiver share an unlimited amount of prior entanglement.

In sharp contrast to the unassisted classical [14] or the quantum capacity [28], the quantity $C_E(\mathcal{E})$ is additive [1]. In particular, it has the single-letter expression $C_E(\mathcal{E}) = \sup_\rho I(\mathcal{E}, \rho)$ in terms of the quantum mutual information

$$I(\mathcal{E}, \rho) = S(\rho) + S(\mathcal{E}(\rho)) - S((\mathcal{E} \otimes I_{A'})(\Psi_{AA'})) =: I(A' : C)_{(\mathcal{E} \otimes I_{A'})(\Psi_{AA'})} ,$$

where $\Psi_{AA'}$ is a purification of the input density operator ρ . (This statement is a generalization [2, 16] of the Holevo-Schumacher-Westmoreland theorem.) The quantity $I(\mathcal{E}, \rho)$ has a number of nice properties: it is positive and concave with respect to the input state ρ .

For channels involving infinite-dimensional state spaces, such capacity results can be adapted by certain limiting procedures – see [17] for a rigorous analysis of the entanglement-assisted capacity. It is necessary to constrain the inputs to obtain meaningful results: one is led to consider the quantity $C_E(\mathcal{E}, \mathbf{N})$ obtained by maximizing over all input states ρ with mean photon number upper bounded by \mathbf{N} , i.e.,

$$C_E(\mathcal{E}, \mathbf{N}) = \sup_{\rho: \langle a^\dagger a \rangle_\rho \leq \mathbf{N}} I(\mathcal{E}, \rho) . \quad (6)$$

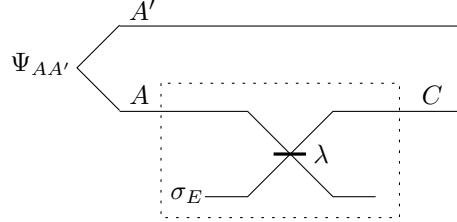


Figure 2: The additive noise channel $\mathcal{E}_{\lambda, \sigma_E}$ with a transmissivity λ -beamsplitter and environment in the state σ_E is schematically illustrated by the dotted box. Its entanglement-assisted capacity $C_E(\mathcal{E}_{\lambda, \sigma_E}, \mathbf{N})$ is obtained by maximizing the mutual information $I(A' : C)$ over all input states ρ_A (with purification $\Psi_{AA'}$) subject to the mean photon number constraint $\text{tr}(a^\dagger a \rho) \leq \mathbf{N}$.

We are interested in estimating (6) when $\mathcal{E} = \mathcal{E}_{\lambda, \sigma_E}$ is an additive noise channel. Such a channel is characterized by the transmissivity $\lambda \in [0, 1]$ and a state σ_E of the environment, see Figure 2. Both the input as well as the environment of the channel $\mathcal{E}_{\lambda, \sigma_E}$ consist of n bosonic modes; the two interact with a beamsplitter of transmissivity λ . The output of the channel is a set of n modes (see Section 8.1 for a precise definition of these expressions). We will focus on $n = 1$ (this often being sufficient because of additivity), although the entropy power inequality also applies for $n > 1$.

The entanglement-assisted capacity of $\mathcal{E}_{\lambda, \sigma_E}$ can easily be upper bounded by twice the maximum output entropy, i.e.,¹

$$C_E(\mathcal{E}_{\lambda, \sigma_E}, \mathbf{N}) \leq 2 \max_{\rho: \langle a^\dagger a \rangle_\rho \leq \mathbf{N}} S(C) \leq 2g(\mathbf{N}_{\max}) . \quad (7)$$

Here the second inequality uses the fact [34] that Gaussian states maximize entropy under a given photon number constraint: $g(x) = (x + 1) \log(x + 1) - x \log x$ is the entropy (in nats) of a Gaussian state with mean photon number x , $\mathbf{N}_{\max} = \lambda\mathbf{N} + (1 - \lambda)\mathbf{N}_E$ is the maximal mean photon number at the output, and $\mathbf{N}_E = \langle a^\dagger a \rangle_{\sigma_E}$ is the mean photon number of the environment. In the limit $\lambda \rightarrow 1$ of perfect transmission, the rhs. of (7) becomes twice the entropy of the input as achievable by dense coding [3]. The following Corollary to our conditional entropy power inequality (Theorem 8.1 below) improves on the upper bound (7).

Corollary 3.1. *Let $\mathcal{E}_{\lambda, \sigma_E}$ be the additive noise channel with transmissivity λ and environment in a state σ_E with mean photon number \mathbf{N}_E . If σ_E is Gaussian, then the entanglement-assisted capacity satisfies*

$$C_E(\mathcal{E}_{\lambda, \sigma_E}, \mathbf{N}) \leq g(\mathbf{N}_{\max}) + \lambda g(\mathbf{N}) - (1 - \lambda)S(E) , \quad (8)$$

where $\mathbf{N}_{\max} = \lambda\mathbf{N} + (1 - \lambda)\mathbf{N}_E$. Moreover, if conjecture (4) holds for all states, then the bound (8) holds even in the case where σ_E is not Gaussian.

¹The proof of (7) follows immediately from (see Figure 2)

$$I(A' : C) = S(C) - S(C|A') = S(C) + S(C|E'D) \leq 2S(C) .$$

Here E' purifies the environment, and D is the second beam-splitter output. The second identity uses the purity of the overall state on $A'CE'D$ and the inequality is subadditivity of the entropy.

We emphasize that Corollary 3.1 is of interest mainly in cases where the channel is not completely characterized as e.g., in quantum cryptography. Under conjecture (4), it gives a universal upper bound independent of the detailed structure of the environment's state σ_E .

Proof. Assume that σ_E is Gaussian such that $\mathcal{E}_{\lambda,\sigma_E}$ is a Gaussian operation. Then the optimization (6) can be restricted to Gaussian states [18]. Consider an arbitrary Gaussian ρ_A , and let $\Psi_{AA'}$ be a Gaussian purification. Then

$$I(A' : C) = S(C) - S(C|A') \leq S(C) - (\lambda S(A|A') + (1 - \lambda)S(E))$$

by the conditional entropy power inequality for Gaussian states. The claim then follows from the maximum entropy principle [34] (i.e., the fact that Gaussian states maximize entropy under a constraint on the second moments) because $S(A|A') = -S(A)$ for a pure state $\Psi_{AA'}$.

The case of a non-Gaussian state σ_E follows in a similar manner (under conjecture (4)). Here we cannot restrict the optimization (6) to Gaussian states. \square

Figure 3 shows a comparison of this bound with the known capacity of the thermal noise channel.

Note that if the state σ_E is Gaussian, then the channel $\mathcal{E}_{\lambda,\sigma_E}$ is a Gaussian operation. Various capacities of such channels have been studied in detail [18] (see [10] for a recent review). In particular, the entanglement-assisted capacity $C_E(\mathcal{E}_{\lambda,\sigma_E}, \mathbf{N})$ is the result of a convex optimization problem: it is obtained by maximizing the mutual information over the (convex) set of covariance matrices M associated with Gaussian states ρ satisfying the photon number constraint². In principle, this can be addressed using efficient numerical algorithms. Furthermore, in special cases, it is possible to give an explicit expression for $C_E(\mathcal{E}_{\lambda,\sigma_E}, \mathbf{N})$: Holevo and Werner [18] computed the entanglement-assisted capacity of the attenuation/amplification channel with classical noise (the most general one-mode channel not involving squeezing). Similarly, the entanglement-assisted capacity of the broadband lossy channel was discussed in detail in [11]. The crucial feature enabling these calculations is gauge-invariance of $\mathcal{E}_{\lambda,\sigma_E}$. Because the entanglement-assisted capacity is additive, this is equivalent to the state σ_E being thermal with respect to the mode operators defining the beam-splitter. In contrast, Corollary 3.1 gives a bound which is applicable to all Gaussian additive channels without further restrictions.

Recall that the entropy power inequality [21] provides an additive upper bound on the classical capacity of thermal noise channels. Because the entanglement-assisted capacity is additive, the conditional entropy power inequality plays a somewhat different role in Corollary 3.1: it substitutes an optimization problem by a bound depending on simple universal parameters (i.e., the entropy and the mean photon number of the environment). One may hope that the conditional entropy power inequality may also be used to address additivity problems in the context of entanglement-assisted communication. A natural candidate problem here is the rate region of the quantum multiple access channel (MAC) characterized by Hsieh, Devetak and Winter [19], where a single-letter formula is not known. The use of conditional entropy power inequalities is especially suggestive in the case of the additive bosonic MAC (see e.g., [35]), where Alice and Bob are connected to a receiver Charlie via two arms of a beamsplitter. As shown by Cezkaj et al. [7] (see also [8]), this scenario exhibits uniquely quantum activation

²This follows from the fact that the maximum mutual information in (6) is achieved on a Gaussian state ρ and the concavity of $I(\mathcal{E}, \rho)$ in ρ , see [18].

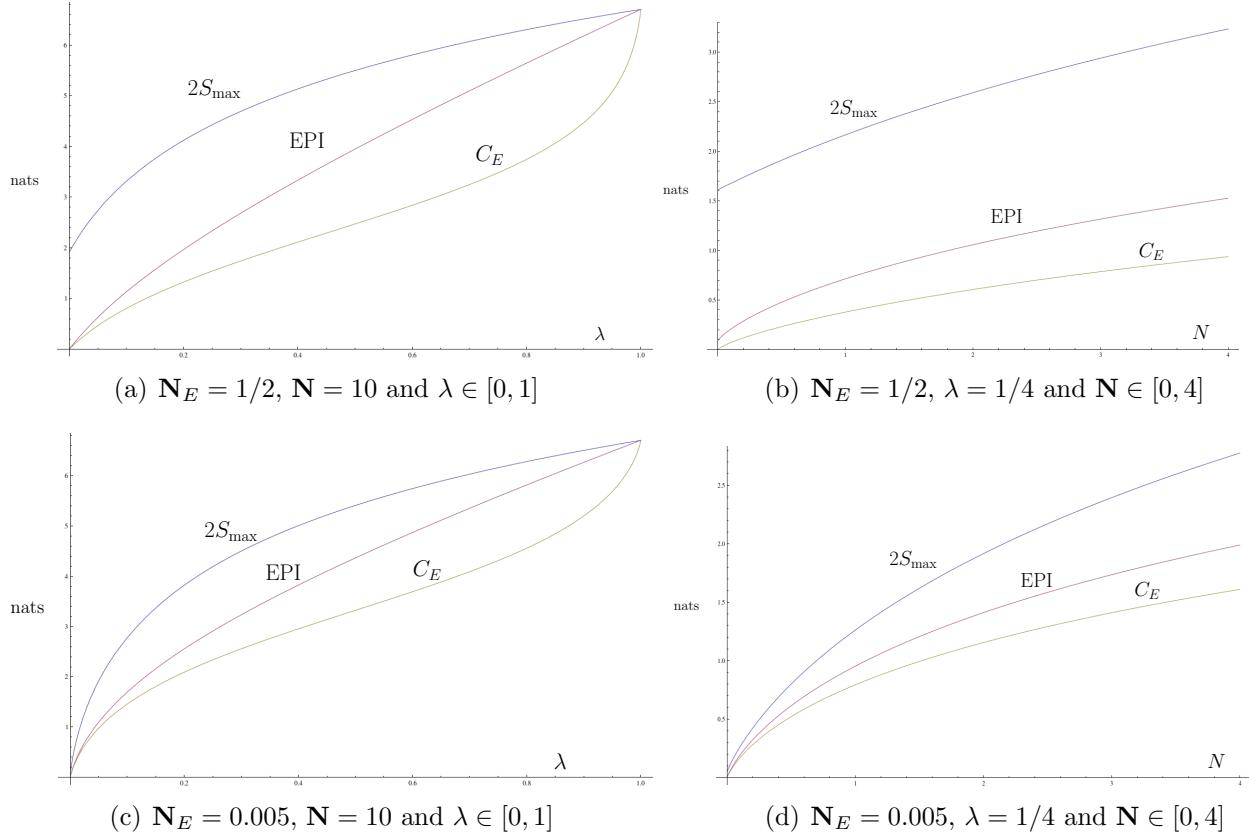


Figure 3: Here we compare the upper bound of Corollary 3.1 with the known capacity of the thermal noise channel (with environment photon number N_E). The latter is given by [18] the expression $C_E(\mathcal{E}_{\lambda, \sigma_E}, N) = g(N) + g(N_{\max}) - g((D + N_{\max} - N - 1)/2) - g((D - N_{\max} + N - 1)/2)$, where $D = \sqrt{(N + N_{\max} + 1)^2 - 4\lambda N(N + 1)}$. We are interested in the case where σ_E is close to the vacuum state corresponding to a pure loss channel. We plot the capacity C_E , the EPI upper bound (8) and the maximum entropy upper bound (7). While in this case, the exact value of the capacity is known, these figures illustrate how the entropy power inequality improves over the trivial bound. Its insensitivity to the exact form of the environment's state may be useful in certain applications. In contrast, the expression for C_E depends on the fact that σ_E is a thermal Gaussian state, that is, the gauge-invariance of $\mathcal{E}_{\lambda, \sigma_E}$.

effects: providing entanglement-assistance to Bob can boost Alice's maximal rate of communication. This is in contrast to analogous classical settings, where providing additional resources to Bob cannot change her maximal rate (the latter is determined by her signal power). While the conditional entropy power inequality can be used to bound the strength of this activation effect, a direct application does unfortunately not appear to yield fundamental new insights into the additivity problem for the bosonic MAC.

Outline

In Section 4, we recall basic definitions. In Section 5, we perturbatively compute the first-order corrections to the symplectic eigenvalues of symmetric matrices relevant in our context. In

Section 6, we apply these perturbative results to obtain the asymptotic scaling of the conditional entropy, as well as its infinitesimal rate of increase when part of a Gaussian state undergoes diffusion. In Section 7, we connect this to Fisher information by establishing a de Bruijin-type identity for conditional entropies. In Section 8, we complete the proof of the entropy power inequality for conditional entropies.

4 Basic definitions

Consider N bosonic modes described by mode operators $\vec{R} = (Q_1, P_1, \dots, Q_N, P_N)$ satisfying the canonical commutation relations

$$[R_k, R_\ell] = iJ_{k,\ell} \quad \text{where} \quad J_N = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{\oplus N}.$$

A Gaussian quantum state $\rho = \rho_{M,\vec{d}}$ on this system is completely specified by its first and second moments, i.e., the displacement vector $\vec{d} = (d_1, \dots, d_{2N}) \in \mathbb{R}^{2N}$ and its (symmetric) covariance matrix M , defined by

$$d_k = \text{tr}(\rho R_k) \quad \text{and} \quad M_{j,k} = \text{tr}(\rho \{R_j - d_j, R_k - d_k\}) .$$

Here $\{A, B\} = AB + BA$. We call $\rho_{M,0}$ a centered Gaussian state and often write ρ_M for it.

A Gaussian operation maps Gaussian states to Gaussian states and is determined by its action on (M, d) . An example is a displacement (or Weyl) operator $W(\vec{\xi})$, $\vec{\xi} \in \mathbb{R}^{2N}$: this is a unitary operation satisfying

$$W(\vec{\xi})\rho_{M,\vec{d}}W(\vec{\xi})^\dagger = \rho_{M,\vec{d}+\vec{\xi}} , \quad (9)$$

for all Gaussian states $\rho_{M,d}$. It will sometimes be convenient to write the conjugation map as $\mathcal{W}_{\vec{\xi}}(\rho) := W(\vec{\xi})\rho W(\vec{\xi})^\dagger$. Statement (9) is equivalent to the Heisenberg action on the mode operators

$$W(\xi)^\dagger R_k W(\xi) = R_k + \xi_k \quad \text{for all } k = 1, \dots, 2N . \quad (10)$$

A matrix S satisfying $SJ_N S^T = J_N$ is called symplectic. A symplectic matrix S uniquely defines a Gaussian unitary U_S by the action on mode operators

$$U_S^\dagger R_j U_S = \sum_{k=1}^{2N} S_{j,k} R_k =: R'_j \quad \text{for all } j = 1, \dots, 2N . \quad (11)$$

Because S is symplectic, the transformed operators $\vec{R}' = (R'_1, \dots, R'_{2N}) = (Q'_1, P'_1, \dots, Q'_N, P'_N)$ again satisfy canonical commutation relations. It is convenient to define the (transformed) creation and annihilation operators

$$a_k^\dagger = \frac{1}{\sqrt{2}}(Q'_k + iP'_k) \quad \text{and} \quad a_k = \frac{1}{\sqrt{2}}(Q'_k - iP'_k) .$$

The associated number operators

$$\hat{n}_k = a_k^\dagger a_k = \frac{1}{2}((Q'_k)^2 + (P'_k)^2) - \frac{1}{2}I, \quad k = 1, \dots, N \quad (12)$$

are mutually commuting, and there is a simultaneous orthonormal eigenbasis of the form $|n\rangle = |n_1, \dots, n_k\rangle$ with $\hat{n}_k|n\rangle = n_k|n\rangle$.

Eq. (11) translates into the action

$$U_S \rho_{M,d} U_S^\dagger = \rho_{SMST,Sd} , \quad (13)$$

on the covariance matrix and the displacement vector. In particular, displacement operators and unitaries of the form U_S can be used to diagonalize any Gaussian state (with (9) and (13)). More precisely, Williamson's theorem [33] states that a symmetric positive definite matrix M can be diagonalized by a symplectic matrix S , i.e.,

$$SMS^T = \text{diag}(\lambda_1, \lambda_1, \lambda_2, \lambda_2, \dots, \lambda_N, \lambda_N) =: D_N(\vec{\lambda}) . \quad (14)$$

We call $\vec{\lambda} = (\lambda_1, \dots, \lambda_N)$ the symplectic eigenvalues of M . Observe that this list may include multiplicities (i.e., $\lambda_i = \lambda_j$ for $i \neq j$). We will occasionally denote the set of distinct symplectic eigenvalues as $\text{Sspec}(M)$. The quantity $\Delta(M) = \min_{\lambda, \tilde{\lambda} \in \text{Sspec}(M), \lambda \neq \tilde{\lambda}} |\lambda - \tilde{\lambda}|$ will be referred to as the symplectic gap of M .

With (13), identity (14) implies that a centered Gaussian state ρ_M with covariance matrix M can be brought into product form as

$$U_S \rho_M U_S^\dagger = \bigotimes_{j=1}^n \frac{e^{-\beta_j \hat{n}_j}}{\text{tr } e^{-\beta_j \hat{n}_j}} , \quad (15)$$

where the inverse temperatures $\beta_j = \beta(\lambda_j)$ are given in terms of the symplectic eigenvalues by

$$\beta(\lambda) = \log \frac{\lambda + 1}{\lambda - 1} . \quad (16)$$

Note that $\beta(\lambda)$ is monotonically decreasing with increasing λ , $\beta(\lambda) < 2$ for $\lambda > 2$, and $\lim_{\lambda \rightarrow \infty} \beta(\lambda) = 0$. If $\lambda_j = 1$, then the factor $\frac{e^{-\beta_j \hat{n}_j}}{\text{tr } e^{-\beta_j \hat{n}_j}}$ in (15) needs to be replaced by the pure 'vacuum' state $|0\rangle\langle 0|_{Q'_j P'_j}$ associated with the mode operators Q'_j, P'_j . Eq. (15) shows that the number states $\{|n\rangle\}_{n \in \mathbb{N}_0^N}$ corresponding to the transformed mode operators \vec{R}' are an eigenbasis of $U_S \rho_M U_S^\dagger$.

The entropy $S(\rho) = -\text{tr}(\rho \log \rho)$ of an N -mode Gaussian state $\rho = \rho_{M,d}$ only depends on the symplectic eigenvalues $\vec{\lambda} = (\lambda_1, \dots, \lambda_N)$ of the covariance matrix. To express it, it is useful to define the mean photon number $\mathbf{N}(\lambda_k)$ in the eigenmode k by the function

$$\mathbf{N}(\lambda) = (\lambda - 1)/2 \quad (17)$$

of a symplectic eigenvalue λ . Then the entropy is given by

$$S(\rho) = \sum_{j=1}^N g(\mathbf{N}(\lambda_j)) \quad \text{with} \quad g(\mathbf{N}) := (\mathbf{N} + 1) \log(\mathbf{N} + 1) - \mathbf{N} \log \mathbf{N} . \quad (18)$$

Let us give a simple bound on the dependence of the entropy on the symplectic eigenvalues.

Lemma 4.1. *Suppose ρ and σ are N -mode Gaussian states with (arbitrarily ordered) symplectic eigenvalues $\vec{\nu} = (\nu_1, \dots, \nu_N)$ and $\vec{\lambda} = (\lambda_1, \dots, \lambda_N)$, respectively, and assume that $\lambda_j \neq 1$ for all $j = 1, \dots, N$. Let $\|\vec{\nu} - \vec{\lambda}\|_\infty = \max_{1 \leq j \leq N} |\nu_j - \lambda_j|$ and $\lambda_* = \min_{1 \leq j \leq N} \lambda_j$. Then*

$$|S(\rho) - S(\sigma)| \leq \frac{N}{2} \left(\|\vec{\nu} - \vec{\lambda}\|_\infty \beta(\lambda_*) + \frac{\|\vec{\nu} - \vec{\lambda}\|_\infty^2}{\lambda_*^2 - 1} \right).$$

The requirement that σ has no symplectic eigenvalue equal to 1 (or equivalently $\lambda_* > 1$) implies that none of the eigenmodes factors out in a pure product state, and we can usually assume this without loss of generality in our considerations below.

Proof. We have $g'(\mathbf{N}) = \log((\mathbf{N} + 1)/\mathbf{N})$, and combining this with (17), we obtain the inverse temperature according to

$$g'(\mathbf{N}(\lambda)) = \beta(\lambda). \quad (19)$$

Furthermore, we have

$$\beta'(\lambda) = -\frac{2}{\lambda^2 - 1}. \quad (20)$$

Let us define the function $G(\lambda) := g(\mathbf{N}(\lambda))$. Because $\mathbf{N}'(\lambda) = \lambda/2$, Eqs. (19) and (20) imply

$$G'(\lambda) = \beta(\lambda)/2 \quad \text{and} \quad G''(\lambda) = -\frac{1}{\lambda^2 - 1}.$$

Hence the Taylor series expansion gives

$$|G(\lambda + \epsilon) - G(\lambda)| \leq \epsilon \beta(\lambda)/2 + \frac{\epsilon^2}{2|\lambda^2 - 1|}.$$

Since β is monotonically decreasing, we get

$$\max_j |G(\nu_j) - G(\lambda_j)| \leq \frac{1}{2} \max_k |\nu_k - \lambda_k| \beta(\min_j \lambda_j) + \frac{\max_k |\nu_k - \lambda_k|^2}{2 \min_j |\lambda_j^2 - 1|}$$

The claim follows from this because $S(\rho) = \sum_{j=1}^N G(\nu_j)$ and $S(\sigma) = \sum_{j=1}^N G(\lambda_j)$. \square

5 Perturbation theory for symplectic eigenvalues

We begin with a straightforward application of degenerate perturbation theory which is illustrative of the method. We will need slightly more involved statements for bipartite systems below (Lemma 5.2). Note that a similar perturbative analysis was used in a different context in [26, Appendix B].

Lemma 5.1 (Perturbation of symplectic spectrum to 0th order). *Let M be a covariance matrix of N modes and consider the symplectic eigenvalues $\vec{\lambda}(\epsilon) = (\lambda_1(\epsilon), \dots, \lambda_N(\epsilon))$ of*

$$M^\infty(\epsilon) = I + \epsilon M,$$

where we assume that $\|M\| = O(1)$ and $\epsilon \ll 1$. Then

$$\lambda_j(\epsilon) = 1 + O(\epsilon) \quad \text{for all } j = 1, \dots, N.$$

Proof. The symplectic eigenvalues $\vec{\lambda}(\epsilon)$ can be obtained from the spectrum $\text{spec}(iJ_N M^\infty(\epsilon))$ of $iJ_N M^\infty(\epsilon)$ since this operator has eigenvalues $(\lambda_1(\epsilon), -\lambda_1(\epsilon), \dots, \lambda_N(\epsilon), -\lambda_N(\epsilon))$. Let us write

$$iJ_N M^\infty(\epsilon) = H + V \quad \text{where} \quad H = iJ_N \quad \text{and} \quad V = i\epsilon J_N M .$$

Observe that the eigenvalues of H are $\{1, -1\}$ and $\|V\| = O(\epsilon)$. This means that we can apply standard degenerate perturbation theory: the spectrum $\text{spec}(H + V)$ is given by

$$\{\mu + \text{spec}(V|_{\mathcal{H}_\mu}) \mid \mu \in \text{spec}(H)\} + O(\epsilon^2) \quad (21)$$

to first order in ϵ , where $V|_{\mathcal{H}_\mu}$ denotes the restriction of V to the degenerate eigenspace \mathcal{H}_μ of H to eigenvalue μ . An immediate consequence is that (by assumption on $\|M\|$), the symplectic eigenvalues of $M^\infty(\epsilon)$ are of the form $\lambda_j(\epsilon) = 1 + O(\epsilon)$ for all $j = 1, \dots, N$. \square

To obtain the correct constant (i.e., the first-order correction in (21)), it is necessary to compute the restriction $V|_{\mathcal{H}_\mu}$ to the eigenspace with eigenvalue μ . For this purpose, it is convenient to use a basis of \mathcal{H}_μ . For example, for the case of Lemma 5.1, we can define the vectors

$$|v^+\rangle = \frac{1}{\sqrt{2}}(i, 1) = \frac{1}{\sqrt{2}}(i|1\rangle + |2\rangle) \quad \text{and} \quad |v^-\rangle = \frac{1}{\sqrt{2}}(1, i) = \frac{1}{\sqrt{2}}(|1\rangle + i|2\rangle)$$

in \mathbb{R}^2 (Here and below, we will often use $|j\rangle$ to denote standard orthonormal basis vectors in \mathbb{R}^{2N} .) Importantly, these vectors satisfy

$$iJ_1|v^\pm\rangle = \pm|v^\pm\rangle , \quad (22)$$

and

$$\langle v^\pm| iJ_1 D_1(\alpha) |v^\pm\rangle = \pm\alpha \quad \langle v^\mp| iJ_1 D_1(\alpha) |v^\pm\rangle = 0 \quad \text{for all } \alpha \in \mathbb{R} . \quad (23)$$

Writing $(0, 0)^{\oplus k} = (0, 0) \oplus \dots \oplus (0, 0)$ (k summands), we can define

$$|v_{j,N}^\pm\rangle = (0, 0)^{\oplus(j-1)} \oplus |v^\pm\rangle \oplus (0, 0)^{\oplus(N-j-1)} .$$

where $|v^\pm\rangle$ is at mode j (of N modes).

Because of (22), the eigenspace \mathcal{H}_1 of H to eigenvalue 1 has orthonormal basis $\{|v_{j,N}^+\rangle\}_{j=1,\dots,N}$. Similarly, the eigenspace \mathcal{H}_{-1} has orthonormal basis $\{|v_{j,N}^-\rangle\}_{j=1,\dots,N}$. The restriction of V to \mathcal{H}_1 is given by the matrix $(\langle v_j^+ | V | v_k^+ \rangle)_{j,k=1,\dots,N}$, and diagonalizing this matrix gives the first-order corrections to the eigenvalue 1 (the degeneracy is lifted). We will omit a more detailed discussion here as we will need a more general version (including a reference system B). However, the proof of the following Lemma proceeds in this fashion; the key here is to apply perturbation theory to a suitably transformed matrix.

Lemma 5.2 (Perturbation of symplectic spectrum to 1st order). *Let*

$$M_{AB} = \begin{pmatrix} M_A & L_{AB} \\ L_{AB}^T & M_B \end{pmatrix}$$

be the covariance matrix of a state on $m + n$ modes with respect to the ordering

$$\vec{R} = (Q_1^A, P_1^A, \dots, Q_m^A, P_m^A, Q_1^B, P_1^B, \dots, Q_n^B, P_n^B)$$

of modes.

(i) Assume that $\|M_{AB}\| = O(1)$ and $\epsilon \ll 1$. Then the symplectic eigenvalues $\vec{\lambda}(\epsilon) = (\lambda_1(\epsilon), \dots, \lambda_{m+n}(\epsilon))$ of the covariance matrix

$$M_{AB}^\infty(\epsilon) = \begin{pmatrix} I_A & 0 \\ 0 & 0 \end{pmatrix} + \epsilon M_{AB} =: I_A \oplus 0_B + \epsilon M_{AB} \quad (24)$$

are of the form

$$\lambda_j(\epsilon) = \eta_j(\epsilon) + O(\epsilon^2) \quad \text{for } j = 1, \dots, m \quad \text{and} \quad \lambda_{m+j}(\epsilon) = \epsilon \nu_j + O(\epsilon^2) \quad \text{for } j = 1, \dots, n$$

where $\vec{\eta}(\epsilon) = (\eta_1(\epsilon), \dots, \eta_m(\epsilon))$ are the eigenvalues of $I_A + \epsilon M_A$ and $\vec{\nu} = (\nu_1, \dots, \nu_n)$ are the eigenvalues of M_B .

(ii) Let $\vec{\lambda} = (\lambda_1, \dots, \lambda_{m+n})$ be the symplectic eigenvalues of M_{AB} and let S_{AB} be the symplectic matrix diagonalizing M_{AB} , i.e.,

$$S_{AB} M_{AB} S_{AB}^T = D_{m+n}(\vec{\lambda}) .$$

Suppose $\epsilon \ll \Delta(M_{AB})$. Let $\vec{\lambda}(\epsilon) = (\lambda_1(\epsilon), \dots, \lambda_{m+n}(\epsilon))$ be the symplectic eigenvalues of

$$M_{AB}^0(\epsilon) = M_{AB} + \epsilon I_A \oplus 0_B . \quad (25)$$

Then degeneracies are split according to

$$|\{\ell \mid \lambda_\ell = \lambda\}| = |\{\ell \mid \lambda_\ell(\epsilon) \in [\lambda - \Delta/2, \lambda + \Delta/2]\}| \quad (26)$$

where ℓ ranges over $\ell \in \{1, \dots, m+n\}$. Furthermore

$$\sum_{\ell: \lambda_\ell(\epsilon) \in [\lambda - \Delta/2, \lambda + \Delta/2]} \lambda_\ell(\epsilon) = \sum_{\ell: \lambda_\ell = \lambda} \left(\lambda + \frac{\epsilon}{2} \text{tr}[S_{AB}(I_A \oplus 0_B)S_{AB}^T]^{(\ell)} \right) + O(\epsilon^2)$$

for each $\lambda \in \text{Sspec}(M_{AB})$. Here $[Z]^{(\ell)}$ denotes the submatrix corresponding to the ℓ -th mode, i.e., the 2×2 -matrix with entries $([Z]^{(\ell)})_{i,j} = Z_{2\ell-1+i, 2\ell-1+j}$, $i, j \in \{0, 1\}$.

Proof of Lemma 5.2 (i)

Our goal is to compute the symplectic eigenvalues $\lambda_1(\epsilon), \dots, \lambda_{m+n}(\epsilon)$ of the covariance matrix

$$M_{AB}^\infty(\epsilon) = I_A \oplus 0_B + \epsilon M_{AB} ,$$

to first order in ϵ . Let $\vec{\nu} = (\nu_1, \dots, \nu_m)$ be the symplectic eigenvalues of M_B , and let S_B be the symplectic matrix diagonalizing M_B , i.e.,

$$S_B M_B S_B^T = D_n(\vec{\nu}) ,$$

where $D_n(\vec{\nu}) = \text{diag}(\nu_1, \nu_1) \oplus \dots \oplus \text{diag}(\nu_n, \nu_n)$. Similarly, let $\vec{\alpha}(\epsilon) = (\alpha_1(\epsilon), \dots, \alpha_m(\epsilon))$ be such that the symplectic eigenvalues of $I_A + \epsilon M_A$ are equal to $(1 + \alpha_1(\epsilon), \dots, 1 + \alpha_m(\epsilon))$. Observe that

$$|\alpha_j(\epsilon)| = O(\epsilon) \quad \text{for } j = 1, \dots, m . \quad (27)$$

Let $S_A(\epsilon)$ be the symplectic matrix diagonalizing $I_A + \epsilon M_A$, i.e.,

$$S_A(\epsilon)(I_A + \epsilon M_A)S_A(\epsilon)^T = I_A + D_m(\vec{\alpha}(\epsilon)) .$$

Finally, let $S_{AB}(\epsilon) = S_A(\epsilon) \oplus S_B$. The symplectic spectrum of $M_{AB}^\infty(\epsilon)$ is identical to that of

$$\hat{M}_{AB}^\infty(\epsilon) := S_{AB}(\epsilon)M_{AB}^\infty(\epsilon)S_{AB}(\epsilon)^T = \begin{pmatrix} I_A + D_m(\vec{\alpha}(\epsilon)) & \epsilon S_A(\epsilon) L_{AB} S_B^T \\ \epsilon S_B L_{AB}^T S_A(\epsilon)^T & \epsilon D_n(\vec{\nu}) \end{pmatrix} .$$

In particular, the eigenvalues of the matrix $iJ_{m+n}\hat{M}_{AB}^\infty(\epsilon)$ are given by

$$\text{spec}(iJ_{m+n}\hat{M}_{AB}^\infty(\epsilon)) = (\lambda_1(\epsilon), -\lambda_1(\epsilon), \dots, \lambda_{m+n}(\epsilon), -\lambda_{m+n}(\epsilon)) . \quad (28)$$

Since $J_{m+n} = J_m \oplus J_n$, we obtain

$$iJ\hat{M}_{AB}^\infty(\epsilon) = H + V \quad \text{where } H = \begin{pmatrix} iJ_m & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} iJ_m D_m(\vec{\alpha}(\epsilon)) & J_m X(\epsilon) \\ J_n X(\epsilon)^T & i\epsilon J_n D_n(\vec{\nu}) \end{pmatrix} ,$$

where we introduced the abbreviation $X(\epsilon) = i\epsilon S_A(\epsilon) L_{AB} S_B^T$. By the assumption $\|M_{AB}\| = O(1)$, we have $\|X(\epsilon)\| = O(\epsilon)$ and $\|D_n(\vec{\nu})\| = O(1)$. Similarly, we have $\|D_m(\vec{\alpha}(\epsilon))\| = O(\epsilon)$ because of Eq. (27). We conclude that $\|V\| = O(\epsilon)$. On the other hand, the eigenvalues of H are $\{0, 1, -1\}$, hence we can apply first-order (degenerate) perturbation theory to compute the spectrum of $iJ\hat{M}_{AB}^\infty(\epsilon)$.

We consider the eigenspaces separately.

- $\mu = 0$: Consider the eigenspace \mathcal{H}_0 of H to eigenvalue $\mu = 0$. It is easy to check that the vectors $\{v_{1,B}^+, v_{1,B}^-, \dots, v_{n,B}^+, v_{n,B}^-\}$ defined by

$$|v_{j,B}^\pm\rangle := 0_A \oplus |v_{j,n}^\pm\rangle$$

are an orthonormal basis of \mathcal{H}_0 . (Here we write 0_A for the zero-vector $(0, 0)^{\oplus m}$.) Since

$$V|v_{k,B}^\tau\rangle = (J_m X|v_{k,n}^\tau\rangle) \oplus (i\epsilon J_n D_n(\vec{\nu})|v_{k,n}^\tau\rangle) ,$$

we get the matrix elements

$$\langle v_{j,B}^\sigma | V | v_{k,B}^\tau \rangle = \langle v_{j,n}^\sigma | i\epsilon J_n D_n(\vec{\nu}) | v_{k,n}^\tau \rangle = \delta_{j,k} \delta_{\sigma,\tau} (\sigma \cdot \epsilon \nu_j)$$

according to (23). In particular, the restriction $V|_{\mathcal{H}_0}$ is described by a diagonal matrix with diagonal elements of the form $\{\pm \epsilon \nu_j\}_{j=1}^n$. With (21), we have found m symplectic eigenvalues of the form

$$\lambda_{m+j}(\epsilon) = \epsilon \nu_j + O(\epsilon^2) \quad \text{for } j = 1, \dots, n . \quad (29)$$

(Only the non-negative entries on the diagonal are symplectic eigenvalues according to (28).) Note that these are simply the eigenvalues of M_B , up to a factor of ϵ .

- $\mu = 1$: An orthonormal basis of \mathcal{H}_1 is given by the vectors $\{|v_{j,A}^+\rangle\}_{j=1}^m$ defined by

$$|v_{j,A}^+\rangle := |v_{j,m}^+\rangle \oplus 0_B$$

where 0_B stands for $(0, 0)^{\oplus n}$. It is easy to check that

$$\langle v_{j,A}^+ | V | v_{k,B}^+ \rangle = \delta_{j,k} \alpha_j(\epsilon) ,$$

i.e., the restriction of V to \mathcal{H}_1 is diagonal when expressed in this basis. In conclusion, we found m symplectic eigenvalues of the form

$$\lambda_j(\epsilon) = 1 + \alpha_j(\epsilon) + O(\epsilon^2) \quad \text{for } j = 1, \dots, m . \quad (30)$$

By definition of $\alpha_j(\epsilon)$, these are equal to the symplectic eigenvalues $\eta_j(\epsilon) = 1 + \alpha_j(\epsilon)$ of $I_A + \epsilon M_A$ in order $O(\epsilon)$.

- $\mu = -1$: Here we find an orthonormal basis with vectors $|v_{j,A}^+\rangle := |v_{j,m}^+\rangle \oplus 0_B$. The restriction of V to \mathcal{H}_{-1} is diagonal with diagonal entries $(-\alpha_1(\epsilon), \dots, -\alpha_m(\epsilon))$. The corresponding eigenvalues are the negatives of (30), consistent with (28).

Proof of Lemma 5.2 (ii)

Let us briefly recall the definitions involved in the statement. We consider the covariance matrix

$$M_{AB}^0(\epsilon) = M_{AB} + \epsilon \begin{pmatrix} I_A & 0 \\ 0 & 0 \end{pmatrix} = M_{AB} + \epsilon I_A \oplus 0_B ,$$

where M_{AB} has symplectic eigenvalues $\vec{\lambda} = (\lambda_1, \dots, \lambda_{m+n})$ and is diagonalized by S_{AB} , i.e.,

$$S_{AB} M_{AB} S_{AB}^T = D_{m+n}(\vec{\lambda}) .$$

We assume that $\epsilon \ll \Delta$, where $\Delta = \Delta(M_{AB})$ is the symplectic gap. Then $M_{AB}^0(\epsilon)$ has the same symplectic eigenvalues as

$$\hat{M}_{AB}^0(\epsilon) = D_{m+n}(\vec{\lambda}) + \epsilon S_{AB}(I_A \oplus 0_B)S_{AB}^T .$$

In particular, it suffices to find the eigenvalues of

$$iJ_{m+n}\hat{M}_{AB}^0(\epsilon) = iJ_{m+n}D_{m+n}(\vec{\lambda}) + i\epsilon J_{m+n}S_{AB}(I_A \oplus 0_B)S_{AB}^T =: H + V .$$

Because $H := iJ_{m+n}D_{m+n}(\vec{\lambda})$ has gap Δ , $\|iJ_{m+n}S_{AB}(I_A \oplus 0_B)S_{AB}^T\| = O(1)$ and our assumption $\epsilon \ll \Delta$, we can apply (degenerate) perturbation theory to compute the spectrum of $iJ\hat{M}_{AB}^0(\epsilon)$.

The operator H has eigenvectors $(w_1^+, \dots, w_{m+n}^+) = (v_{1,A}^+, \dots, v_{m,A}^+, v_{1,B}^+, \dots, v_{n,B}^+)$ with eigenvalues $\lambda_1, \dots, \lambda_{m+n}$. In particular, for $\lambda \in \text{Sspec}(M_{AB})$, the eigenspace \mathcal{H}_λ of H is

$$\mathcal{H}_\lambda = \text{span}\{|w_\ell^+\rangle \mid 1 \leq \ell \leq m+n \text{ with } \lambda_\ell = \lambda\} .$$

Suppose that restriction $V|_{\mathcal{H}_\lambda}$ of $V := i\epsilon J_{m+n}S_{AB}(I_A \oplus 0_B)S_{AB}^T$ to \mathcal{H}_λ has eigenvalues $\theta_1, \dots, \theta_{\dim \mathcal{H}_\lambda}$. According to (degenerate) first-order perturbation theory, the matrix $iJ_{m+n}\hat{M}_{AB}^0(\epsilon)$ has $\dim \mathcal{H}_\lambda$ eigenvalues of the form

$$\lambda + \theta_j + O(\epsilon^2) \quad \text{for } j = 1, \dots, \dim \mathcal{H}_\lambda , \quad (31)$$

where $\theta_j = O(\epsilon)$. Furthermore, the list (31) includes all eigenvalues $\lambda_\ell(\epsilon)$ of $iJ_{m+n}\hat{M}_{AB}^0(\epsilon)$ in the interval $[\lambda - \Delta/2, \lambda + \Delta/2]$ and the number of such eigenvalues is equal to (26) by the assumption $\epsilon \ll \Delta$. We conclude that

$$\begin{aligned} \sum_{\ell: \lambda_\ell(\epsilon) \in [\lambda - \Delta/2, \lambda + \Delta/2]} \lambda_\ell(\epsilon) &= (\dim \mathcal{H}_\lambda) \cdot \lambda + \sum_{j=1}^{\dim \mathcal{H}_\lambda} \theta_j + O(\epsilon^2) \\ &= (\dim \mathcal{H}_\lambda) \cdot \lambda + \text{tr}(V|_{\mathcal{H}_\lambda}) + O(\epsilon^2) \end{aligned} \quad (32)$$

Because the restriction $V|_{\mathcal{H}}$ can be expressed as

$$V|_{\mathcal{H}_\lambda} = \sum_{j,k: \lambda_j = \lambda_k = \lambda} \langle w_j^+ | V | w_k^+ \rangle \cdot |w_j^+ \rangle \langle w_k^+| ,$$

we obtain

$$\text{tr}(V|_{\mathcal{H}_\lambda}) = i\epsilon \sum_{\ell: \lambda_\ell = \lambda} \langle w_\ell^+ | J_{m+n} S_{AB} (I_A \oplus 0_B) S_{AB}^T | w_\ell^+ \rangle . \quad (33)$$

Combining (33) with $\dim \mathcal{H}_\lambda = |\{\ell : \lambda_\ell = \lambda\}|$ and Eq. (32) gives

$$\sum_{\ell: \lambda_\ell(\epsilon) \in [\lambda - \Delta/2, \lambda + \Delta/2]} \lambda_\ell(\epsilon) = \sum_{\ell: \lambda_\ell = \lambda} (\lambda + i\epsilon \langle w_\ell^+ | J_{m+n} S_{AB} (I_A \oplus 0_B) S_{AB}^T | w_\ell^+ \rangle) + O(\epsilon^2) . \quad (34)$$

Since $|w_\ell^+ \rangle = \frac{1}{\sqrt{2}}(i|2\ell-1\rangle + |2\ell\rangle)$ and $J_{m+n} = \sum_{k=1}^{m+n} |2k-1\rangle \langle 2k| - |2k\rangle \langle 2k-1|$, we have $\langle w_\ell^+ | J_{m+n} = \frac{1}{\sqrt{2}}(i\langle 2\ell | + \langle 2\ell-1 |)$, and this takes the form

$$\langle w_\ell^+ | J_{m+n} S_{AB} (I_A \oplus 0_B) S_{AB}^T | w_\ell^+ \rangle = \frac{1}{2} (iZ_{2\ell, 2\ell-1} - iZ_{2\ell-1, 2\ell} + Z_{2\ell, 2\ell} + Z_{2\ell-1, 2\ell-1})$$

where $Z = S_{AB} (I_A \oplus 0_B) S_{AB}^T$. Since $Z^T = Z$ is symmetric, this is equal to

$$\langle w_\ell^+ | J_{m+n} S_{AB} (I_A \oplus 0_B) S_{AB}^T | w_\ell^+ \rangle = \frac{1}{2} \text{tr}[S_{AB} (I_A \oplus 0_B) S_{AB}^T]^{(\ell)} , \quad (35)$$

and the claim follows by inserting (35) into (34).

6 Conditional entropy and diffusion

A central tool in the proof of the quantum entropy power inequality [20] is the one-parameter semigroup $\{e^{t\mathcal{L}}\}_{t \geq 0}$ of completely positive trace-preserving maps generated by the ‘diffusion’ Liouvillean \mathcal{L} (see [20] for a definition of the latter). For all $t \geq 0$, the map $e^{t\mathcal{L}}$ is Gaussian and can therefore be defined in terms of its action on the covariance matrix M and the displacement vector \vec{d} : A Gaussian state ρ described by (M, \vec{d}) is transformed into a Gaussian state $\rho(t) = e^{t\mathcal{L}}(\rho)$ with covariance matrix $(M(t) = M + tI, \vec{d})$, i.e., the transformation governing the evolution for time t is

$$\begin{aligned} \rho &\mapsto e^{t\mathcal{L}}(\rho) \\ (M, \vec{d}) &\xrightarrow{e^{t\mathcal{L}}} (M + tI, \vec{d}) \end{aligned} \quad \text{for all covariance matrices } M \text{ and displacement vectors } \vec{d} .$$

In this section, we revisit and extend statements of [20] about the behavior of the entropy $S(e^{t\mathcal{L}}(\rho))$ as a function of time t . In particular, we specialize to Gaussian initial states ρ and extend our considerations to conditional entropies.

More precisely, we consider the case where diffusion acts only on a subset of modes. Concretely, assume that our system is bipartite, with system A consisting of m modes, and system B consisting of n modes. Diffusion for time t acting on the modes in A only is described by the superoperator $e^{t\mathcal{L}_A} \otimes I_B$ (where I_B is the identity superoperator on B). In particular, this family of superoperators is specified by the transformation

$$\begin{array}{ccc} \rho_{AB} & \mapsto & (e^{t\mathcal{L}_A} \otimes I_B)(\rho_{AB}) =: \rho_{AB}(t) \\ (M_{AB}, \vec{d}_{AB}) & \xrightarrow{e^{t\mathcal{L}_A} \otimes I_B} & (M_{AB} + t(I_A \oplus 0_B), \vec{d}_{AB}) \end{array} \quad (36)$$

for all covariance matrices $M_{AB} \in \text{Mat}_{2(m+n)}(\mathbb{R})$ and displacement vectors $\vec{d}_{AB} \in \mathbb{R}^{2(m+n)}$. We will examine the conditional entropy $S(A|B)_{\rho_{AB}(t)}$ for the evolved state $\rho_{AB}(t)$, given some Gaussian initial state $\rho_{AB} = \rho_{AB}(0)$. In Section 6.1, we show that $S(A|B)_{\rho_{AB}(t)}$ scales as a universal function for $t \rightarrow \infty$ (independent of the initial state ρ_{AB}). In Section 6.2, we derive an explicit expression for the infinitesimal rate of change of this quantity in terms of the covariance matrix of ρ_{AB} .

6.1 Scaling of the conditional entropy in the infinite-time limit

In the infinite-time-limit, the entropy of the time-evolved state $e^{t\mathcal{L}}(\rho)$ scales as a universal function of time t which is independent of the initial state ρ . This statement was shown for general states in [20, Corollary 3.4]; here we give a simple argument for Gaussian states for completeness. We will also need this statement in the proof of Lemma 6.2 which deals with conditional entropies.

Lemma 6.1 (Scaling of entropy in the infinite-time limit under diffusion). *Let ρ be an (arbitrary) Gaussian state of N modes. Then*

$$\lim_{t \rightarrow \infty} |S(e^{t\mathcal{L}}(\rho)) - N \cdot g((t-1)/2)| = 0 .$$

Proof. Let M be the covariance matrix of ρ . The covariance matrix of $\rho(t) = e^{t\mathcal{L}}(\rho)$ has the form $M + tI = tM^\infty(1/t)$. Therefore, the symplectic eigenvalues are of the form $\nu_j = t(1 + O(1/t)) = t + O(1)$ according to Lemma 5.1. For $t \geq 1$, the matrix tI is a valid covariance matrix with symplectic eigenvalues $\lambda_j = t$ for $j = 1, \dots, N$. Let $\sigma(t)$ be the centered Gaussian state with covariance matrix tI . Lemma 4.1 applied to $\rho(t)$ and $\sigma(t)$ gives

$$|S(\rho(t)) - S(\sigma(t))| \leq \frac{N}{2} O(1) \left(\beta(t) + \frac{1}{t^2 - 1} \right) \rightarrow 0 \quad \text{for } t \rightarrow \infty .$$

Since $S(\sigma(t)) = N \cdot g(\mathbf{N}(t)) = N \cdot g((t-1)/2)$, the claim follows. \square

A similar statement holds for conditional entropies:

Lemma 6.2 (Scaling of conditional entropy in the infinite-time limit under diffusion). *Let ρ_{AB} be a Gaussian state of $(m+n)$ modes A and B . Define $\rho_{AB}(t) = (e^{t\mathcal{L}_A} \otimes I_B)(\rho_{AB})$. Then*

$$\lim_{t \rightarrow \infty} |S(A|B)_{\rho_{AB}(t)} - m \cdot g((t-1)/2)| = 0 .$$

Comparing Lemma 6.2 with Lemma 6.1 suggests that $(e^{t\mathcal{L}_A} \otimes I_B)(\rho_{AB})$ approaches a product state for large times. Quantifying this convergence for general (possibly non-Gaussian) states may provide an avenue to proving conjecture 4.

Proof. Let $\vec{\nu} = (\nu_1, \dots, \nu_n)$ be the symplectic eigenvalues of the covariance matrix M_B . Note that we can assume without loss of generality that

$$\nu_j \neq 1 \quad \text{for all } j = 1, \dots, n. \quad (37)$$

Indeed, if we had $\nu_j = 1$ for some $j \in \{1, \dots, n\}$, then the corresponding eigenmode in B is in a pure state and the state ρ_{AB} factorizes. This is true also for the time-evolved state $\rho_{AB}(t)$ (because B is unaffected by the evolution), hence such eigenmodes do not contribute to the entropy $S(A|B)_{\rho(t)}$ and can be traced out.

The covariance matrix of the state $\rho_{AB}(t)$ can be written in the form

$$M_{AB} + tI_A \oplus 0_B = t(I_A \oplus 0_B + 1/t \cdot M_{AB}) =: tM_{AB}^\infty(1/t).$$

Applying Lemma 5.2 (i) with $\epsilon = 1/t$ (and multiplying the resulting eigenvalues by t), we conclude that its symplectic eigenvalues are

$$\lambda_j = \eta_j(t) + O(1/t) \quad \text{for } j = 1, \dots, m \quad \text{and} \quad \lambda_{m+j} = \nu_j + O(1/t) \quad \text{for } j = 1, \dots, n,$$

where $\{\eta_j(t)\}$ are the symplectic eigenvalues of $tI_A + M_A$. This means that up to order $O(1/t)$, the symplectic spectrum of $\rho_{AB}(t)$ is identical to the one associated with the product state $e^{t\mathcal{L}}(\rho_A) \otimes \rho_B$. In summary, we have the list of symplectic eigenvalues

$\rho_{AB}(t)$	$e^{t\mathcal{L}_A}(\rho_A)$	ρ_B
$\underbrace{(\vec{\eta}(t) + O(1/t), \vec{\nu}(t) + O(1/t))}_{=: \tilde{\eta}}$	$\vec{\eta}(t)$	$\vec{\nu}(t)$

Let $\tilde{\rho}_A$ be a Gaussian state of m modes with symplectic spectrum $\tilde{\eta}$, and let $\tilde{\rho}_B$ be a Gaussian state of n modes with symplectic spectrum $\tilde{\nu}$. Then

$$S(\rho_{AB}(t)) = S(\tilde{\rho}_A) + S(\tilde{\rho}_B). \quad (38)$$

We can apply Lemma 4.1 to the states $\tilde{\rho}_A$ and $e^{t\mathcal{L}_A}(\rho_A)$. Since the latter has symplectic eigenvalues $\vec{\eta}(t) = (\eta_1(t), \dots, \eta_m(t))$, we get

$$|S(\tilde{\rho}_A) - S(e^{t\mathcal{L}}(\rho_A))| \leq \frac{m}{2} \left(O(1/t) \beta(\eta_*(t)) + O(1/t^2) \frac{1}{\eta_*(t)^2 - 1} \right) \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

Here we used the fact that $\eta_j(t) = t + O(1)$ for all $1 \leq j \leq m$, hence $\eta_*(t) = \min_j \eta_j(t) \rightarrow \infty$ for $t \rightarrow \infty$. Similarly, applying Lemma 4.1 to the states $\tilde{\rho}_B$ and ρ_B (and using assumption (37)) gives

$$|S(\tilde{\rho}_B) - S(\rho_B)| \leq \frac{n}{2} \left(O(1/t) \cdot \beta(\nu_*) + O(1/t^2) \frac{1}{\min_j |\nu_j^2 - 1|} \right) \rightarrow 0 \quad \text{for } t \rightarrow \infty.$$

Inserting these upper bounds into (38) and using the triangle inequality then gives

$$|S((e^{t\mathcal{L}_A} \otimes I_B)(\rho_{AB})) - S(e^{t\mathcal{L}_A}(\rho_A) \otimes \rho_B)| \rightarrow 0 \quad \text{for } t \rightarrow \infty.$$

Because the states $\rho_{AB}(t) = (e^{t\mathcal{L}_A} \otimes I_B)(\rho_{AB})$ and $\sigma_{AB}(t) = e^{t\mathcal{L}_A}(\rho_A) \otimes \rho_B$ have the same reduced density operator $\rho_B(t) = \sigma_B(t) = \rho_B$, the previous statement implies that the difference between their conditional entropies also vanishes in the limit, that is,

$$|S(A|B)_{\rho_{AB}(t)} - S(A|B)_{\sigma(t)}| \rightarrow 0 \quad \text{for } t \rightarrow \infty. \quad (39)$$

Since $\sigma_{AB}(t)$ is a product state, we have

$$S(A|B)_{\sigma(t)} = S(\sigma_A(t)). \quad (40)$$

The claim then follows from the triangle inequality, i.e.,

$$|S(A|B)_{\rho_{AB}(t)} - m \cdot g((t-1)/2)| \leq |S(A|B)_{\rho_{AB}(t)} - S(\sigma_A(t))| + |S(\sigma_A(t)) - m \cdot g((t-1)/2)|$$

because the first term on the rhs. goes to 0 for $t \rightarrow \infty$ according to (39) and (40), whereas the second term goes to 0 according to Lemma 6.1. \square

6.2 Rate of increase of the conditional entropy

Next we compute the infinitesimal rate of increase of the conditional entropy under the process (36).

Lemma 6.3 (Rate of conditional entropy increase under diffusion). *Consider bipartite system AB of $m+n$ modes. Let ρ_{AB} be a Gaussian quantum state whose covariance matrix M_{AB} has symplectic eigenvalues $\vec{\lambda} = (\lambda_1, \dots, \lambda_{m+n})$. Let S_{AB} be a symplectic matrix such that $S_{AB}M_{AB}S_{AB}^T = D_{m+n}(\vec{\lambda})$. Define $\rho_{AB}(t) = (e^{t\mathcal{L}_A} \otimes I_B)(\rho_{AB})$. Then*

$$\frac{d}{dt} \Big|_{t=0} S(A|B)_{\rho_{AB}(t)} = \frac{1}{4} \sum_{\ell=1}^{m+n} \beta(\lambda_\ell) \text{tr} \left([S_{AB}(I_A \oplus 0_B)S_{AB}^T]^{(\ell)} \right),$$

where $[Z]^{(\ell)}$ is the 2×2 submatrix corresponding to the ℓ -th mode, $([Z]^{(\ell)})_{i,j} = Z_{2\ell-1+i, 2\ell-1+j}$ for $i, j \in \{0, 1\}$.

Proof. Because the reduced density operator $\text{tr}_A \rho_{AB}(t) = \rho_B$ is independent of time, we have

$$\frac{d}{dt} \Big|_{t=0} S(A|B)_{\rho_{AB}(t)} = \frac{d}{dt} \Big|_{t=0} S(AB)_{\rho_{AB}(t)}. \quad (41)$$

To evaluate the rate of change of $S(AB)_{\rho_{AB}(t)}$, let $(\lambda_1(t), \dots, \lambda_{m+n}(t))$ be the symplectic eigenvalues of $\rho_{AB}(t)$. According to expression (18) for the entropy, we have

$$S(\rho_{AB}(t)) = \sum_{j=1}^{m+n} g(\mathbf{N}(\lambda_j(t))) = \sum_{\lambda \in \text{Sspec}(M_{AB})} \sum_{\ell: \lambda_\ell(t) \in [\lambda - \frac{\Delta}{2}, \lambda + \frac{\Delta}{2}]} g(\mathbf{N}(\lambda_\ell(t))) \quad (42)$$

where we reexpressed the summation using the symplectic gap $\Delta = \Delta(M_{AB})$. With the expression (17) for the mean photon number, Eq. (19) and the chain rule for differentiation, we have

$$g(\mathbf{N}(\lambda(t))) = g(\mathbf{N}(\lambda(0))) + \frac{t\beta(\lambda(0))}{2} \cdot \lambda'(0) + O(t^2)$$

Observe that $\rho_{AB}(t)$ has covariance matrix $M_{AB}^0(t)$ and we can restrict our attention to times $t \ll \Delta(M_{AB})$ without loss of generality. Hence we can apply Lemma 5.2 (ii). We obtain

$$\sum_{\ell: \lambda_\ell(t) \in [\lambda - \frac{\Delta}{2}, \lambda + \frac{\Delta}{2}]} g(\mathbf{N}(\lambda_\ell(t))) = g(\mathbf{N}(\lambda)) \cdot |\{\ell \mid \lambda_\ell = \lambda\}| + \frac{t\beta(\lambda)}{4} \sum_{\ell: \lambda_\ell = \lambda} \text{tr}[S_{AB}(I_A \oplus 0_B)S_{AB}^T]^{(\ell)} + O(t^2) .$$

for any $\lambda \in \text{Sspec}(M_{AB})$. Taking the derivative at $t = 0$ and inserting into (42) therefore gives

$$\frac{d}{dt} \Big|_{t=0} S(\rho_{AB}(t)) = \frac{1}{4} \sum_{\lambda \in \text{Sspec}(M_{AB})} \beta(\lambda) \sum_{\ell: \lambda_\ell = \lambda} \text{tr}[S_{AB}(I_A \oplus 0_B)S_{AB}^T]^{(\ell)} ,$$

which is the claim because of (41). \square

7 Diffusion, translations and Fisher information

A key element in the proof of the classical entropy power inequality is de Bruijin's identity; it relates the infinitesimal rate of entropy increase to the Fisher information of a family of translated distributions. In [20], a quantum version of this statement in terms of the diffusion semigroup and phase space translations was given. Here we derive a generalization of this statement for conditional entropies (but specialized to Gaussian states). Our proof proceeds by direct calculation and does not involve any technical subtleties associated with formal computations involving infinite-dimensional systems.

We begin by recalling the relevant definitions. Consider a one-parameter family $\{\rho^{(\theta)}\}_{\theta \in \mathbb{R}}$ of states depending smoothly on the parameter θ . The divergence-based Fisher information of this family (at $\theta_0 \in \mathbb{R}$) is defined as the quantity

$$J(\rho^{(\theta)}; \theta)|_{\theta=\theta_0} = \frac{d^2}{d\theta^2} S(\rho^{(0)} \| \rho^{(\theta)})|_{\theta=\theta_0} ,$$

where $S(\rho \| \sigma) = \text{tr}(\rho \log \rho - \rho \log \sigma)$ is the relative entropy or divergence. Two straightforward but important consequences of this definition are the reparametrization identities

$$J(\rho^{(c\theta)}; \theta)|_{\theta=0} = c^2 J(\rho^{(\theta)}; \theta)|_{\theta=0} \quad \text{and} \quad J(\rho^{(\theta+c)}; \theta)|_{\theta=\theta_0} = J(\rho^{(\theta+c)}; \theta)|_{\theta=\theta_0+c} \quad (43)$$

for $c \in \mathbb{R}$ and its additivity

$$J(\rho_A^{(\theta)} \otimes \rho_B^{(\theta)}; \theta)|_{\theta=\theta_0} = J(\rho_A^{(\theta)}; \theta)|_{\theta=\theta_0} + J(\rho_B^{(\theta)}; \theta)|_{\theta=\theta_0} . \quad (44)$$

The latter follows from the additivity of the relative entropy under tensor products. Furthermore, because of the monotonicity $S(\mathcal{E}(\rho) \| \mathcal{E}(\sigma)) \leq S(\rho \| \sigma)$ of the relative entropy under CPTP maps \mathcal{E} the Fisher information also satisfies monotonicity (see [20]), i.e.,

$$J(\mathcal{E}(\rho^{(\theta)}); \theta)|_{\theta=0} \leq J(\rho^{(\theta)}; \theta)|_{\theta=0} . \quad (45)$$

de Bruijin's identity involves the family $\{\rho^{(\theta, R_k)} = \mathcal{W}_{\theta|k}(\rho)\}_{\theta \in \mathbb{R}}$ of states obtained by translating an N -mode state ρ in the direction R_k in phase space, where $k \in \{1, \dots, 2N\}$. Let us write

$$J(\rho) := \sum_{k=1}^{2N} J(\rho^{(\theta, R_k)}; \theta)|_{\theta=0} \quad (46)$$

for the sum of the corresponding Fisher informations. de Bruijin's identity (shown in [20]) relates this to the rate of entropy increase under diffusion, i.e.,

$$\frac{d}{dt} \Big|_{t=0} S(e^{t\mathcal{L}}(\rho)) = \frac{1}{4} J(\rho) . \quad (47)$$

In this section, we derive a version of (47) for Gaussian states which involves an auxiliary system: it quantifies the rate of increase in the conditional entropy $S(A|B)$ when A undergoes diffusion. In contrast to [20], the proof given here proceeds by direct computation. In Section 7.1, we compute the Fisher information of a family of states obtained by translating a Gaussian in phase space. In Section 7.2, we combine this with Lemma 6.3 to prove the de Bruijin identity.

7.1 Conditional Fisher information of translated Gaussian states

Let $\rho_{M,\vec{d}}$ denote an N -mode Gaussian state with covariance matrix M and displacement $\vec{d} \in \mathbb{R}^{2N}$. Suppose S is a symplectic matrix such that $SMS^T = D(\vec{\gamma})$ is diagonal. Let $\vec{\theta} \in \mathbb{R}^{2N}$ be arbitrary. We will need the following formula for the relative entropy of $\rho_{M,\vec{d}}$ and a displaced state $\rho_{M,\vec{d}+\vec{\theta}}$: we have

$$S(\rho_{M,\vec{d}}\|\rho_{M,\vec{d}+\vec{\theta}}) = F(\vec{\gamma}) + \frac{1}{2} \sum_{j=1}^N \beta(\gamma_j) ((S\theta)_{2j-1}^2 + (S\theta)_{2j}^2) . \quad (48)$$

where F is a function of the symplectic eigenvalues only.

Proof. By the invariance $S(U\rho U^\dagger\|U\sigma U^\dagger) = S(\rho\|\sigma)$ of the relative entropy under unitaries, and applying displacement operators as well as the unitary U_S , we have

$$S(\rho_{M,\vec{d}}\|\rho_{M,\vec{d}+\vec{\theta}}) = S(\rho_{M,0}\|\rho_{M,\vec{\theta}}) = S(\rho_{M,-\vec{\theta}}\|\rho_{M,0}) = S(\rho_{D(\gamma),-\vec{\theta}}\|\rho_{D(\gamma),0}) .$$

It hence suffices to analyze $S(\rho_{D(\gamma),\vec{\eta}}\|\rho_{D(\gamma),0})$, where $\vec{\eta} = -S\vec{\theta} \in \mathbb{R}^{2N}$. Because

$$\rho_{D(\gamma),\vec{\eta}} = \rho_{D_1(\gamma_1),(\eta_1,\eta_2)} \otimes \cdots \otimes \rho_{D_1(\gamma_N),(\eta_{2N-1},\eta_{2N})}$$

is a product state, we obtain (using the additivity of the relative entropy for product states)

$$S(\rho_{M,d}\|\rho_{M,\vec{d}+\vec{\theta}}) = S(\rho_{D(\gamma),\vec{\eta}}\|\rho_{D(\gamma),0}) = \sum_{j=1}^N S(\rho_{D_1(\gamma_j),(\eta_{2j-1},\eta_{2j})}\|\rho_{D_1(\gamma_j),0}) .$$

The claim therefore follows from Lemma 7.1. □

Lemma 7.1. *Let $\vec{\eta} = (\eta_Q, \eta_P) \in \mathbb{R}^2$, $\gamma \in \mathbb{R}$, and let $D = \text{diag}(\gamma, \gamma)$ be the covariance matrix of a single mode Gaussian state. Then*

$$S(\rho_{D,\vec{\eta}}\|\rho_{D,0}) = \frac{\beta}{2}(\mathbf{N} + \eta_Q^2 + \eta_P^2) - g(\mathbf{N}) - \log(1 - e^{-\beta}) , \quad (49)$$

where the mean photon number \mathbf{N} and the inverse temperature β are given by (17) and (16), respectively.

Proof. For brevity, let us write $\rho_D = \rho_{D,0}$ for the centered state. Then we have

$$S(\rho_{D,\vec{\eta}}\|\rho_D) = -S(\rho_D) - \text{tr}(\rho_{D,\vec{\eta}} \log \rho_{D,0}) = -g(\mathbf{N}(\gamma)) - \text{tr}(\rho_{D,\vec{\eta}} \log \rho_{D,0}) . \quad (50)$$

To compute the latter term, we use the expression $\rho_D = \frac{e^{-\beta\hat{n}}}{\text{tr}(e^{-\beta\hat{n}})} = (1 - e^{-\beta})e^{-\beta\hat{n}}$, where $\hat{n} = a^\dagger a = \frac{1}{2}(Q^2 + P^2 - 1)$ is the number operator and $\beta = \beta(\gamma) = \log(\gamma+1)/(\gamma-1)$ the inverse temperature. By taking the logarithm, one gets

$$-\text{tr}(\rho_{D,\vec{\eta}} \log \rho_D) = -\log(1 - e^{-\beta}) + \beta \text{tr}(\rho_{D,\vec{\eta}} \hat{n}) .$$

Using the fact that $\rho_{D,\vec{\eta}} = W(\vec{\eta})\rho_D W(\vec{\eta})^\dagger$ for the Weyl operator $W(\vec{\eta})$ and the fact that

$$W(\vec{\eta})^\dagger \hat{n} W(\vec{\eta}) = \frac{1}{2}((Q + \eta_Q)^2 + (P + \eta_P)^2 - 1)$$

according to (10) and Definition (12), we get

$$\begin{aligned} -\text{tr}(\rho_{D,\vec{\eta}} \log \rho_D) &= -\log(1 - e^{-\beta}) + \frac{\beta}{2} \text{tr}(\rho_D((Q + \eta_Q)^2 + (P + \eta_P)^2 - 1/2)) \\ &= -\log(1 - e^{-\beta}) + \frac{\beta}{2} (\text{tr}(\rho_D \hat{n}) + \eta_Q^2 + \eta_P^2) . \end{aligned}$$

In the last line, we made use of the fact that ρ_D is centered. The claim follows by combining this with (50). \square

With (48), we can easily compute the Fisher information of a family of displaced states.

Lemma 7.2 (Fisher information of displaced states). *Let M_{AB} be the covariance matrix of a centered state $\rho_{M,0}$ of $m+n$ modes, where $S_{AB}M_{AB}S_{AB}^T = D_{m+n}(\vec{\gamma})$. Fix some $k \in \{1, \dots, 2m\}$ and consider the family of states $\{\rho^{(\theta, R_k)}\}_{\theta \in \mathbb{R}}$,*

$$\rho^{(\theta, R_k)} = \rho_{M_{AB}, \theta|k\rangle} = \mathcal{W}_{\theta|k\rangle}(\rho_{M_{AB}, 0}) \quad (51)$$

obtained by displacing the state $\rho_{M,0}$ in the direction R_k by an amount $\theta \in \mathbb{R}$. Then

$$J(\rho^{(\theta, R_k)}; \theta)|_{\theta=0} = \sum_{j=1}^{m+n} \beta(\gamma_j)(S_{2j-1, k}^2 + S_{2j, k}^2) .$$

Proof. With (48), we obtain

$$S(\rho_{M,0}\|\rho^{(\theta, R_k)}) = F(\gamma) + \frac{\theta^2}{2} \sum_{j=1}^{m+n} \beta(\gamma_j)(S_{2j-1, k}^2 + S_{2j, k}^2) .$$

The Fisher information is the second derivative of this quantity with respect to θ at $\theta = 0$, hence the claim follows. \square

7.2 The de Bruijin identity for conditional entropies of Gaussian states

It will be convenient to define the conditional Fisher information

$$J(A|B)_{\rho_{AB}} := \sum_{k=1}^{2m} J(\rho^{(\theta, R_k)}; \theta)|_{\theta=0} \quad (52)$$

by summing over the modes corresponding to system A only. Observe that many properties of the Fisher information carry over to this definition: for example, we have monotonicity

$$J(\mathcal{E}(A)|B) \leq J(A|B), \quad (53)$$

for any CPTPM acting on \mathcal{E} , where these quantities are evaluated on the states ρ_{AB} and $(\mathcal{E} \otimes I_B)(\rho_{AB})$, respectively.

By combining Lemma 6.3 and Lemma 7.2, we obtain a proof of the following statement.

Theorem 7.3 (de Bruijin identity for Gaussian states and conditional entropy). *Let ρ_{AB} be a centered Gaussian state of $m+n$ modes. Define $\rho_{AB}(t) = (e^{t\mathcal{L}_A} \otimes I_B)(\rho_{AB})$. Then*

$$\frac{d}{dt} \Big|_{t=0} S(A|B)_{\rho_{AB}(t)} = \frac{1}{4} J(A|B)_{\rho_{AB}}.$$

Proof. It is straightforward to check that $S(I_A \oplus 0_B)S^T$ has diagonal elements of the form $(S(I_A \oplus 0_B)S^T)_{\ell,\ell} = \sum_{k=1}^{2m} S_{\ell,k}^2$. Hence

$$\begin{aligned} \frac{d}{dt} \Big|_{t=0} S((e^{t\mathcal{L}_A} \otimes I_B)(\rho_{AB})) &= \frac{1}{4} \sum_{j=1}^{m+n} \beta(\gamma_j) ((S(I_A \oplus 0_B)S^T)_{2j-1,2j-1} + (S(I_A \oplus 0_B)S^T)_{2j,2j}) \\ &= \frac{1}{4} \sum_{j=1}^{m+n} \sum_{k=1}^{2m} \beta(\gamma_j) (S_{2j-1,k}^2 + S_{2j,k}^2) \end{aligned}$$

according to Lemma 6.3. We conclude from Lemma 7.2 and Definition (52) that

$$\frac{d}{dt} \Big|_{t=0} S((e^{t\mathcal{L}_A} \otimes I_B)(\rho_{AB})) = \frac{1}{4} J(A|B)_{\rho_{AB}}$$

The claim then follows because $\rho_B(t) = \rho_B$ does not depend on time (cf. (41)). \square

8 The entropy power inequality for conditional entropy

Having established the de Bruijin identity for conditional entropies as well as the asymptotic scaling of the conditional entropies under diffusion, it is straightforward to prove the entropy power inequality for conditional entropies. Indeed, this follows the pattern of known classical proofs [9], with minor modifications because we are considering conditional entropies. It relies heavily on the Fisher information inequality (a consequence of data processing, as shown by Zamir [36]).

We introduce the necessary definitions in Section 8.1. The conditional Fisher information inequality and the conditional entropy power inequality are derived subsequently in Sections 8.2 and 8.3.

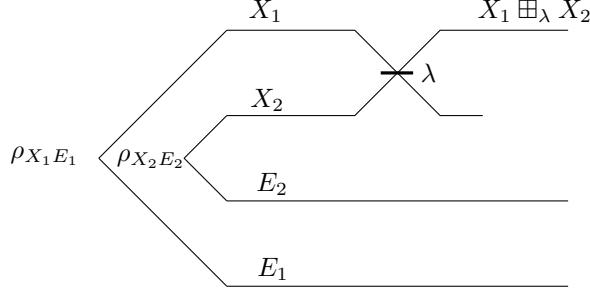


Figure 4: This quantum circuit defines the states $\rho_{X_1 \boxplus_\lambda X_2}$ and $\rho_{(X_1 \boxplus_\lambda X_2) E_1 E_2}$ for any product input $\rho_{X_1 E_1} \otimes \rho_{X_2 E_2}$.

8.1 Beam splitters, product states and auxiliary systems

Consider two systems X_j , $j = 1, 2$ with N modes each and associated mode operators $\{Q_k^{(j)}, P_k^{(j)}\}_{k=1}^N$. A beam-splitter with transmissivity $\lambda \in [0, 1]$ acting on $X_1 X_2$ is the Gaussian unitary U_{S_λ} described by the symplectic matrix

$$S_\lambda = \begin{pmatrix} \sqrt{\lambda} I_{2N} & \sqrt{1-\lambda} I_{2N} \\ \sqrt{1-\lambda} I_{2N} & -\sqrt{\lambda} I_{2N} \end{pmatrix} \quad (54)$$

with respect to the ordering $(Q_1^{(1)}, P_1^{(1)}, \dots, Q_N^{(1)}, P_N^{(1)}, Q_1^{(2)}, P_1^{(2)}, \dots, Q_N^{(2)}, P_N^{(2)})$ of modes. We are interested in the beam-splitter map $\mathcal{E}_\lambda = \mathcal{E}_\lambda^{X_1 X_2 \mapsto Y}$, which is obtained by letting X_1, X_2 interact according to U_{S_λ} , and discarding the second set of N modes. That is, it is a map from $2N$ input modes to N output modes; we call the latter Y . Formally, the map \mathcal{E}_λ is defined as

$$\mathcal{E}_\lambda(\rho_{X_1 X_1}) = \text{tr}_{X_2} \left(U_\lambda \rho_{X_1 X_2} U_\lambda^\dagger \right) ,$$

where tr_{X_2} denotes the partial trace of over the second set X_2 of modes. We will denote the output system (i.e., the set of modes X_1 at the end of this process) by Y , i.e., think of $\mathcal{E}_\lambda = \mathcal{E}_\lambda^{X_1 X_2 \mapsto Y}$ as a map from systems $X_1 X_2$ to an output system Y (of N modes). Since the partial trace tr_{X_2} is a Gaussian map, the map \mathcal{E}_λ is Gaussian and completely determined by its action on covariance matrices and displacement vectors. This is

$$\left(\begin{pmatrix} M_{X_1} & \rho_{X_1 X_2} \\ L_{X_1 X_2}^T & M_{X_2} \end{pmatrix}, (\vec{d}_{X_1}, \vec{d}_{X_2}) \right) \xrightarrow{\mathcal{E}_\lambda} (M_Y, \vec{d}_Y) \quad \mapsto \mathcal{E}_\lambda(\rho_{X_1 X_2})$$

where

$$\begin{aligned} M_Y &= \lambda M_{X_1} + (1 - \lambda) M_{X_2} + \sqrt{\lambda(1 - \lambda)}(L_{X_1 X_2} + L_{X_1 X_2}^T) \\ \vec{d}_Y &= \sqrt{\lambda} \vec{d}_{X_1} + \sqrt{1 - \lambda} \vec{d}_{X_2} . \end{aligned} \quad (55)$$

This follows immediately from (54).

We will consider input states that are products (across the bipartition $X_1 : X_2$). It will be convenient to introduce the following maps and states as summarized in Figure 4:

the state $\rho_{X_1 \boxplus_\lambda X_2}$, given ρ_{X_1}, ρ_{X_2} : For product inputs $\rho_{X_1} \otimes \rho_{X_2}$, the transformation (55) specializes to

$$\begin{aligned} \rho_{X_1} \otimes \rho_{X_2} &\mapsto \sigma_Y := \mathcal{E}_\lambda^{X_1 X_2 \mapsto Y}(\rho_{X_1} \otimes \rho_{X_2}) \\ \left(\begin{pmatrix} M_{X_1} & 0 \\ 0 & M_{X_2} \end{pmatrix}, (d_{X_1}, d_{X_2}) \right) &\xrightarrow{\mathcal{E}_\lambda} (\lambda M_{X_1} + (1 - \lambda) M_{X_2}, \sqrt{\lambda} \vec{d}_{X_1} + \sqrt{1 - \lambda} \vec{d}_{X_2}) , \end{aligned}$$

where we assume that ρ_{X_j} is a Gaussian state described by (M_{X_j}, \vec{d}_{X_j}) , $j = 1, 2$. We will denote the output state σ_Y obtained in this fashion by $\rho_{X_1 \boxplus_\lambda X_2}$.

the state $\rho_{(X_1 \boxplus_\lambda X_2) E_1 E_2}$ given $\rho_{X_1 E_1}$ and $\rho_{X_2 E_2}$: More generally, consider the map $I_{E_1 E_2} \otimes \mathcal{E}_\lambda^{X_1 X_2 \mapsto Y}$, where E_j , $j = 1, 2$ are auxiliary systems with L modes each. We will only need a description of its action on product states $\rho_{X_1 E_1} \otimes \rho_{X_2 E_2}$, where $\rho_{X_j E_j}$ are Gaussian states with covariance matrix and displacement vectors

$$M_{X_j E_j} = \begin{pmatrix} M_{X_j} & L_{X_j E_j} \\ L_{X_j E_j}^T & M_{E_j} \end{pmatrix} \quad \vec{d}_j = (\vec{d}_{X_j}, \vec{d}_{E_j}) \quad \text{for } j = 1, 2 . \quad (56)$$

It is straightforward to verify that this is given by

$$\begin{aligned} \rho_{X_1 E_1} \otimes \rho_{X_2 E_2} &\mapsto \sigma_{Y E_1 E_2} := (I_{E_1 E_2} \otimes \mathcal{E}_\lambda^{X_1 X_2 \mapsto Y})(\rho_{X_1 E_1} \otimes \rho_{X_2 E_2}) \\ (M_{X_1 E_1} \oplus M_{X_2 E_2}, (\vec{d}_1, \vec{d}_2)) &\mapsto (M_{Y E_1 E_2}, (\vec{d}_Y, \vec{d}_{E_1}, \vec{d}_{E_2})) \end{aligned}$$

where

$$\begin{aligned} M_{Y E_1 E_2} &= \begin{pmatrix} \lambda M_{X_1} + (1 - \lambda) M_{X_2} & \sqrt{\lambda} L_{X_1 E_1} & \sqrt{1 - \lambda} L_{X_2 E_2} \\ \sqrt{\lambda} L_{X_1 E_1}^T & M_{E_1} & 0 \\ \sqrt{1 - \lambda} L_{X_2 E_2}^T & 0 & M_{E_2} \end{pmatrix} \\ \vec{d}_Y &= \sqrt{\lambda} \vec{d}_{X_1} + \sqrt{1 - \lambda} \vec{d}_{X_2} . \end{aligned} \quad (57)$$

8.2 The conditional Fisher information inequality for beamsplitters

In [20, Lemmas 3.2 and 6.1], it was shown that the beam-splitter map is compatible with both diffusion and translations in the following sense. For all $t \geq 0$, $w_1, w_2 \in \mathbb{R}$, $\vec{\theta} \in \mathbb{R}^{2m}$, we have the following identities of Gaussian maps:

$$\mathcal{E}_\lambda^{X_1 X_2 \mapsto Y} \circ (e^{t \mathcal{L}_{X_1}} \otimes e^{t \mathcal{L}_{X_2}}) = e^{t \mathcal{L}_Y} \circ \mathcal{E}_\lambda^{X_1 X_2} \quad (58)$$

$$\mathcal{E}_\lambda^{X_1 X_2 \mapsto Y} \circ (\mathcal{W}_{w_1 \vec{\theta}}^{X_1} \otimes \mathcal{W}_{w_2 \vec{\theta}}^{X_2}) = \mathcal{W}_{w \vec{\theta}}^Y \circ \mathcal{E}_\lambda^{X_1 X_2 \mapsto Y} \quad \text{where } w = \sqrt{\lambda} w_1 + \sqrt{1 - \lambda} w_2 . \quad (59)$$

This was then used to show that the quantity J (cf. (46)) satisfies the Fisher information inequality

$$J(X_1 \boxplus_\lambda X_2) \leq \lambda J(X_1) + (1 - \lambda) J(X_2) . \quad (60)$$

The proof of (61) follows immediately from the monotonicity (45) of the divergence-based Fisher information, its additivity (44), the reparametrization identity (43), as well as the compatibility properties (58), (59). We will omit the corresponding argument here; it was discovered in the classical context by Zamir [36].

Here we argue briefly that the quantity (52) satisfies an analogous inequality, that is,

$$J(X_1 \boxplus_{\lambda} X_2 | E_1 E_2) \leq \lambda J(X_1 | E_1) + (1 - \lambda) J(X_2 | E_2) . \quad (61)$$

Indeed, it is clear that the identities (58) and (59) still hold if we replace the maps $e^{t\mathcal{L}_Z}$, $\mathcal{E}_{\lambda}^{X_1 X_2 \mapsto Y}$ and \mathcal{W}_{ξ}^Z by their ‘stabilized’ versions (obtained by adjoining an identity)

$$\begin{array}{ccc} e^{t\mathcal{L}} & \mapsto & e^{t\mathcal{L}} \otimes I_{E_1 E_2} \\ \mathcal{E}_{\lambda}^{X_1 X_2 \mapsto Y} & \mapsto & \mathcal{E}_{\lambda}^{X_1 X_2 \mapsto Y} \otimes I_{E_1 E_2} \\ \mathcal{W}_{\xi}^Z & \mapsto & \mathcal{W}_{\xi}^Z \otimes I_{E_1 E_2} . \end{array}$$

Furthermore, the quantity $J(A|B)_{\rho_{AB}}$ is also monotonic (cf. (53)). Because each of the terms $J(\rho^{\theta, R_k}; \theta)|_{\theta=0}$ constituting $J(A|B)_{\rho_{AB}}$ is additive and satisfies the reparametrization identities, we can apply Zamir’s proof again (carrying along $E_1 E_2$) and obtain the conditional Fisher information inequality (61).

8.3 The conditional entropy power inequality for Gaussian states

The entropy power inequality we prove relates the conditional entropy $S(Y|E_1 E_2)_{\sigma}$ of the state

$$\sigma_{YE_1 E_2} = (\mathcal{E}_{\lambda} \otimes I_{E_1 E_2})(\rho_{X_1 E_1} \otimes \rho_{X_2 E_2}) := \rho_{(X_1 \boxplus_{\lambda} X_2) E_1 E_2}$$

to the conditional entropies $S(X_j|E_j)$ of the two (Gaussian) states $\rho_{X_j E_j}$, $j = 1, 2$.

Theorem 8.1 (Conditional entropy power inequality for Gaussian states). *Let $\rho_{X_1 E_1}$ and $\rho_{X_2 E_2}$ be arbitrary Gaussian states. Then*

$$S(X_1 \boxplus_{\lambda} X_2 | E_1 E_2) \geq \lambda S(X_1 | E_1) + (1 - \lambda) S(X_2 | E_2) .$$

Note that the proof outlined here combined with the discussion in [20] (respectively Stam’s proof [29]) should also provide the inequality

$$e^{S(X_1 \boxplus_{\lambda} X_2 | E_1 E_2)} \geq \frac{1}{2} e^{S(X_1 | E_1) / n} + \frac{1}{2} e^{S(X_2 | E_2) / n}$$

where X_1 and X_2 have n modes. We do not discuss this version here for brevity.

Proof. Let the covariance matrices and displacement vectors of $\rho_{X_j E_j}$, $j = 1, 2$, be given by (56). The corresponding covariance matrices are

$$M_{X_j(t) E_j} = \begin{pmatrix} M_{X_j(t)} & L_{X_j E_j} \\ L_{X_j E_j}^T & M_{E_j} \end{pmatrix} \quad \text{where } M_{X_j(t)} = M_{X_j} + tI .$$

For $t \geq 0$, define the function

$$\delta(t) := S(X_1(t) \boxplus_{\lambda} X_2(t) | E_1 E_2) - \lambda S(X_1(t) | E_1) - (1 - \lambda) S(X_2(t) | E_2) ,$$

where the entropies are evaluated on the states $\rho_{X_j(t) E_j}$, $j = 1, 2$ and the result $\rho_{(X_1(t) \boxplus_{\lambda} X_2(t)) E_1 E_2}$ of letting these interact with the beamsplitter (as discussed in Section 8.1). According to the ‘stabilized’ version of (58), we have

$$\begin{aligned} \delta(t) &= S((X_1 \boxplus X_2)(t) | E_1 E_2) - \lambda S(X_1(t) | E_1) - (1 - \lambda) S(X_2(t) | E_2) \\ &= S(Y(t) | E_1 E_2) - \lambda S(X_1(t) | E_1) - (1 - \lambda) S(X_2(t) | E_2) . \end{aligned}$$

This shows that $\delta(t)$ is the difference of conditional entropies of time-evolved states for different initial states $\rho_{YE_1E_2}$, $\rho_{X_1E_1}$ and $\rho_{X_2E_2}$ at $t = 0$. We conclude with Lemma 6.2 that

$$\lim_{t \rightarrow \infty} \delta(t) = 0. \quad (62)$$

On the other hand, we have according to the de Bruijn identity (Theorem 7.3)

$$4\delta'(t) = J(X_1(t) \boxplus_{\lambda} X_2(t) | E_1 E_2) - \lambda J(X_1(t) | E_1) - (1 - \lambda) J(X_2(t) | E_2).$$

This identity, together with Fisher information inequality (61) imply that $\delta'(t) \leq 0$ for all $t \geq 0$. With (62), this shows that $\delta(0) \geq 0$, which is the claim. \square

Acknowledgements

I would like to thank the organizers of the workshop ‘Beyond iid in quantum information theory’. I also thank Reinhard Werner, Graeme Smith and Jon Yard for discussions, and gratefully acknowledge support by NSERC.

References

- [1] C. Adami and N. J. Cerf. von Neumann capacity of noisy quantum channels. *Phys. Rev. A*, 56:3470–3483, Nov 1997.
- [2] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Trans. Inf. Theory*, 48:2637–2655, 2002.
- [3] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69:2881–2884, Nov 1992.
- [4] P. Bergmans. A simple converse for broadcast channels with additive white Gaussian noise. *Information Theory, IEEE Transactions on*, 20(2):279–280, March 1974.
- [5] N. Blachman. The convolution inequality for entropy powers. *Information Theory, IEEE Transactions on*, 11(2):267 – 271, apr 1965.
- [6] M. H. M. Costa and T. M. Cover. On the similarity of the entropy power inequality and the Brunn-Minkowski inequality. *IEEE Transactions on Information Theory*, 30(6):837–839, 1984.
- [7] L. Czekaj, J. K. Korbicz, R. W. Chhajlany, and P. Horodecki. Quantum superadditivity in linear optics networks: Sending bits via multiple-access gaussian channels. *Phys. Rev. A*, 82:020302, Aug 2010.
- [8] L. Czekaj, J. K. Korbicz, R. W. Chhajlany, and P. Horodecki. Schemes of transmission of classical information via quantum channels with many senders: Discrete- and continuous-variable cases. *Phys. Rev. A*, 85:012316, Jan 2012.

- [9] A. Dembo, T.M. Cover, and J.A. Thomas. Information theoretic inequalities. *Information Theory, IEEE Transactions on*, 37(6):1501–1518, nov 1991.
- [10] J. Eisert and M.M. Wolf. Gaussian quantum channels. In *Quantum Information with Continuous Variables of Atoms and Light*, London, 2007. Imperial College Press. arXiv:quant-ph/0505151.
- [11] V. Giovannetti, S. Lloyd, L. Maccone, and P. W. Shor. Entanglement assisted capacity of the broadband lossy channel. *Phys. Rev. Lett.*, 91:047901, Jul 2003.
- [12] S. Guha. *Multiple-User Quantum Information Theory for Optical Communication Channels*. PhD thesis, Massachusetts Institute of Technology, June 2008.
- [13] S. Guha, B. I. Erkmen, and J. H. Shapiro. The entropy photon-number inequality and its consequences. 2007.
- [14] M. Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5:255–257, 2009.
- [15] P. Hayden, R. Jozsa, D. Petz, and A. Winter. Structure of states which satisfy strong subadditivity of quantum entropy with equality. *Communications in Mathematical Physics*, 246(2):359–374, April 2004.
- [16] A. S. Holevo. On entanglement-assisted classical capacity. *J. Math. Phys.*, 43(4326), 2002.
- [17] A. S. Holevo. Entanglement-assisted capacity of constrained channels. In *First International Symposium on Quantum Informatics, Proc. SPIE 5128*, volume 62, July 2003.
- [18] A. S. Holevo and R. F. Werner. Evaluating capacities of bosonic gaussian channels. *Phys. Rev. A*, 63:032312, Feb 2001.
- [19] M. Hsieh, I. Devetak, and A. Winter. Entanglement-assisted capacity of quantum multiple-access channels. *Information Theory, IEEE Transactions on*, 54(7):3078–3090, July.
- [20] R. König and G. Smith. The entropy power inequality for quantum systems, 2012. arXiv:1205.3409.
- [21] R. König and G. Smith. Limits on classical communication from quantum entropy power inequalities. *Nature Photonics*, 7(254):142–146, February 2013.
- [22] E. H. Lieb. Proof of an entropy conjecture of Wehrl. *Comm. Math. Phys.*, 62(1):35–41, 1978.
- [23] T. Liu and P. Viswanath. An extremal inequality motivated by multiterminal information-theoretic problems. *Information Theory, IEEE Transactions on*, 53(5):1839–1851, May 2007.
- [24] Y. Oohama. The rate-distortion function for the quadratic Gaussian CEO problem. *Information Theory, IEEE Transactions on*, 44(3):1057–1070, May 1998.

- [25] L. Ozarow. On a source coding problem with two channels and three receivers. *The Bell Syst. Tech. J.*, 59:1909–1921, December 1980.
- [26] A. Serafini, J. Eisert, and M. M. Wolf. Multiplicativity of maximal output purities of Gaussian channels under Gaussian inputs. *Phys. Rev. A*, 71:012320, Jan 2005.
- [27] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, October 1948.
- [28] G. Smith and J. Yard. Quantum communication with zero-capacity channels. *Science*, 321(5897):1812–1815, 2008.
- [29] A. J. Stam. Some inequalities satisfied by the quantities of information of Fisher and shannon. *Information and Control*, 2(2):101 – 112, 1959.
- [30] S. J. Szarek and D. Voiculescu. Volumes of restricted Minkowski sums and the free analogue of the entropy power inequality. *Communications in Mathematical Physics*, 178(3):563–570, 1996.
- [31] G. Toscani. An information-theoretic proof of Nash’s inequality, June 2012.
- [32] S. Verdu and D. Guo. A simple proof of the entropy-power inequality. *Information Theory, IEEE Transactions on*, 52(5):2165 –2166, may 2006.
- [33] J. Williamson. On the algebraic problem concerning the normal forms of linear dynamical systems. *American Journal of Mathematics*, 58(1):pp. 141–163, 1936.
- [34] M. M. Wolf, G. Giedke, and J. I. Cirac. Extremality of Gaussian quantum states. *Phys. Rev. Lett.*, 96:080502, Mar 2006.
- [35] B. J. Yen and J. H. Shapiro. Multiple-access bosonic communications. *Phys. Rev. A*, 72:062312, Dec 2005.
- [36] R. Zamir. A proof of the Fisher information inequality via a data processing argument. *Information Theory, IEEE Transactions on*, 44(3):1246 –1250, may 1998.